A MULTIMODAL CONSUMER INFORMATION SERVER WITH IVR MENU

Martin Damhuis, Marc Peeters, Louis Boves
Royal PTT Netherlands N.V., PTT Research
St. Paulusstraat 4, 2264XZ Leidschendam
The Netherlands

ABSTRACT

This paper describes the development of a fully automatic multimodal information system for the consumer market. The system will be able to provide information on a large number of topics via a single telephone number. The eventual system will integrate Interactive Voice Response, speech recognition, speaker verification, Direct Dial In, Calling Line Identification, facsimile and electronic mail. The present version is limited to DTMF input and voice and facsimile output. The architecture of the system described in this paper allows successive addition of other technologies.

I. INTRODUCTION

A consumer Information Server must be suitable for use by the general public. This means that the system must fulfil certain demands:

- Information should be interesting for a large group of potential users.
- This server should add value to existing services.
- Services should be easy to understand and simple to use.
- All services should be fully automatic.
- All services should have a uniform and user friendly interface.

The present version of the information server that we are developing employs an IVR menu with DTMF input as the means by which the callers can interact with the system and selectively retrieve information items from a database. Two applications are described in detail, viz. tele-shopping for consumer telecommunication products and a service giving information about consumer tests of telecommunication hardware. Both services use speech synthesis and facsimile to deliver the requested information to the caller. The information in the system can be updated remotely, partially by means of a PC and a network, partly via another IVR application that allows the system administrator to record new voice prompts and to add new information to the system's database. The following sections deal with the system configuration, a description of the information server and the services that are already implemented; section 4 describes a number of developments which are planned for the near future: the use of advanced techniques like Calling Line Identification (CLI), speech recognition, speaker

verification and text to speech synthesis will allow us to improve the information server.

II. SYSTEM CONFIGURATION

2.1 System architecture

The system architecture of the multimodal consumer information server is based on the concept that is visualised in Figure 1.

Figure 1: concept of multimodal information server

This concept distinuishes between a number of abstract platforms that are needed to give a consumer access to information server:

- The input platform makes it possible to order information.
- The output platform makes it possible to receive and perceive information.
- The network platform makes it possible to transmit information from one point to an other.
- The server platform provides information.

A service description, made according to this concept, is independent of the chosen platform. The input and output platform are not necessarily identical. For example, a consumer orders the information via an interactive voice response menu and the information is delivered via electronic mail. In this example the telephone is input platform and the PC is output platform. This makes the consumer information service multimodal.

The system architecture of the multimodal consumer information server is given in Figure 2.
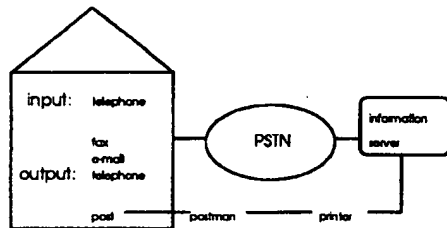
Fig 2: system architecture of multimodal information server

The application according to the system architecture can provide a fully automatic service end to end.

2.1.1 Input platform

The input platform of the present version of the system is a telephone which is used to access, select and perhaps order information via an IVR menu, over the Public Switched Telephone Network (PSTN). Telephone access to an automatic information server has several advantages: telephones are always easy to find, even when a customer is away from home or office. If telephone input is limited to recognition of DTMF codes, it is a very unattractive medium for inputting large amounts of information. Yet, complex applications may need much input. For this reason, we envisage the use of speech recognition in the information ordering process instead of or in addition to the keypad of the telephone with DTMF tones.

The output during the interactive voice response menu is of course restricted to speech (or audio) because of the telephone.

2.1.2 Ouput platform

There is a range of output platforms; the selection depends on which of the platforms available to the consumer is most appropriate:

- The telephone can be used as output platform, so the information is given via the same platform at the input. This is useful if the message is short.
- Post as output platform is, like the telephone, a platform accessible to everybody. The post is a powerful platform to receive information, but it needs some time to deliver and the delivery address is not easy to provide via an IVR menu with DTMF input.
- Electronic mail is a way to deliver information automatically and in several different multimedia formats. However, few consumers have access to electronic mail and there is also a problem for the consumer to provide the mail-address via an IVR menu with DTMF input.
- The fax is an output platform that is useful to deliver text and graphics. This way of information delivery can benefit from increasing popularity of the consumer fax. The fax number of the consumer is easy to provide in an IVR menu with DTMF input.

To implement the output platform in our system an existing platform is used that provides multimodal delivery via fax and mail, viz. Europub [1]. At the end of a successful query menu an electronic mail message is send to the Europub platform, which takes care of the delivery in the mode requested by the caller.

2.2 System Hardware

The workstation used for realising the input-platform is based on a PC running the OS/2 operating system, a Rhetorex Voice Card and driver software. The Voice Card has a four channel analogue telephone interface. The platform uses system messages encoded in a 4-bit ADPCM format. It can be connected to other cards via an ABC connector compatible with the MVIP interface. Other technologies can be added to the system by connecting facsimile cards and speech recognition boards to the voice card via this MVIP interface.

2.3 Service Creation Tools

For creating speech interactive services it is important to have a software tool that enables a service developer to create custom applications rapidly and which additionally provides system management functions. In this project we used the Show N Tel application generator which is a 4GL tool with a point and click icon based graphical user interface. Applications can be developed quickly and emerging standards with regards to menu interaction can be incorporated into the tool, allowing services to have a consistent "look and feel". Together with the Rhetorex software this application generator software includes tools for service development, voice prompt recording and editing tools, database functionality and support, management information and maintenance tools.

2.4 Use of PTT Telecom Style Guide

An important aspect of designing IVR menus is that all services offered by PTT Telecom should have the same 'look and feel'. IVR menus are not always easy to design and it has been found that users quickly are discouraged from using systems that are difficult to understand or tedious to use. A style guide is essentially a set of rules that will help developers to create user friendly IVR menus. A consistent style includes using standard phrasing and arrangement of the menu structure to ensure that the consumer will find learning how to use a new system much easier. The applications that are part of the Information Server have been built in conformance with the style guide adopted by PTT Telecom. This style guide for menus has been developed at PTT Research. It is partly based on previous work of standard bodies such as the European Telecommunication Standards Institute (ETSI), and associations such as the Voice Messaging User Interface Forum (VMUIF). The final style guide was made by evaluating human factors issues of existing automated telephone services in the Netherlands and by performing usability tests of prototypes of new voice applications (menus with DTMF or speech recognition) that were developed at PTT Research. The features of the PTT Telecom style guide that are most important for the

development of the information server are summarised in Table 1.

Table 1: Specific features of PTT Telecom style guide

Key	Functions
1	menu choice, yes
2	menu choice, no
7	go back (only used when * is not possible)
9	main menu
0	request help, operator fallback, entrance test
*	go back, abort entry, repeat
#	go ahead, terminate entry, interrupt
type ahead must be allowed at menus and data entry	
each option is followed by its action number	
menu structure: max. four choices and three layers deep (except for sequential entry of data)	

III. OVERVIEW INFORMATION SERVER

There is a broad range of applications that can be offered with the system configuration. In this section two promising applications are described that fulfil the demands that are mentioned in the introduction. Eventually, the applications can be accessed via three alternative charging methods:

1. Two or more separate telephone numbers, each with an appropriate charge.
2. One telephone number that guides the caller through an IVR menu, where the choice can be made between different applications with appropriate tariffs. Presently, this service is not yet available in the Netherlands.
3. The Scopecard number[1] which can be provided via DTMF input and that allows charging the caller for the cost made accessing the information service.

The following paragraphs describe two applications that are implemented for the Dutch consumer market. The two applications can be accessed via one telephone number (alternative 2) and the charging is done via the Scopecard number. The only relationship that these applications have is that they are offered by one company.

3.1 Teleshopping service for the consumer telecommunications products

PTT Telecom offers two ways of buying products:
- Via shops in major cities.
- Via human telephone operators.

A new way of selling products is fully automatic teleshopping. By providing consumers with brochures containing information about the portfolio, product ordering numbers and a special telephone number a teleshopping service for the consumer telecommunications products is created. In the automatic service three different steps have to be taken in the IVR menu to order some products:
- Ordering
- Identification
- Order verification

Ordering
In the IVR menu the caller can get an overview of the products that are offered (which can be send to him by fax or post); moreover, one has the opportunity to order products. To order the products the ordering numbers must be provided by DTMF input. The IVR menu gives the caller the opportunity to order several products or to end the ordering process. At the end of the ordering process the caller hears the list of ordered products and the total costs. If the total cost is below a given threshold, the system will advise the caller that the order is rejected, unless additional products are ordered.

Identification
Identification of the caller is done via the Scopecard. The Scopecard number and a PIN has to be entered. The payment of the bill can be also done with this unique number.

Order verification
If the caller wants a conformation of the order he/she can indicate this at the end of the IVR menu; one can also select the way in which the conformation should be forwarded (mail, post or fax). DTMF tones are used to identify the fax number. In the future the Scopecard number can be used to the identify post address, the mail address or the fax number of the caller. The information of the caller and his/her order is put into a data-file which is sent to a remote system via electronic mail. This file contains information like a personal ID (the Scopecard number), the way the information has to be delivered (fax, electronic mail or paper mail), and an ID of the ordered products or requested information (e.g. the list with the overview of the products that can be ordered).

3.2 Consumers Hardware Test Information Service
Consumers often have the need for an independent opinion about certain products they want to purchase. For example you want to buy a (new) cordless telephone set, something you do once every 10 years, and you want the latest test of an independent organisation. The consumer information server can be the key to the latest test of cordless telephone sets and provide you with the information via voice response, facsimile or (post/electronic) mail.

In the IVR menu two different steps have to be taken to find and order the information:
- Identification of the information
- Information retrieval identification

Identification of the information
The IVR menu will first of all give an overview of the Consumers Hardware Test Information that is available. Due to the length of the list a skip and scan is implemented. The caller can also ask for a list of the available tests, which can be sent to the caller by fax or post.

To allow the caller to go faster through the long list of items; a skip and scan method is implemented. Via the IVR menu the caller can find his way through the information desired. The application will read a short conclusion of the selected test report. The caller can then decide to stop or request more detailed information about the test.

Information retrieval identification

If the consumer does want more detailed information about the tests he/she must indicate this and select the way in which the information should be sent (mail, post or fax). DTMF tones are used to identify the fax number. In the future the Scopecard number will be used to identify post address, the electronic mail address or the fax number of the consumer directly.

2.3 Remote update.

An information or service provider will require some means of updating the information on the system. This can be done locally on the actual system but it is often more convenient to have the ability to do this remotely. Some of the information in the system can be updated remotely with a another IVR application that allows the system administrator to record new prompts. To update the information that has to be sent by e.g. facsimile, an update service by means of a PC and modem or PC via LAN is a possibility. This allows documents to be sent in a variety of formats, e.g. ASCII text, PostScript or one of the standard PC formats e.g. facsimile format.

IV. FUTURE DEVELOPMENTS

Probably the most attractive feature of the advanced Information Server under development is that it provides access to a rich set of services via a single telephone number. However, the advantage of having to remember only a single number creates the disadvantage of a more complex service selection. Thus, there is a need to replace the selection menu with a mixed-initiative dialogue. Advanced features of the telecommunication network can also be used to enhance the service quality. The system architecture, based on a number of abstract platforms, allows easy integration of new technologies.

4.1 Speech Recognition

Work is under way to offer automatic speech recognition (ASR) as an additional input platform. In its first version ASR will emulate DTMF recognition, making the service accessible from rotary dial phones. The next step will improve menu navigation by allowing the caller to speak the names of services and products instead of giving numerical codes. Eventually we envisage the replacement of the menu interface with a mixed initiative dialogue interface, using the speech and NLP technology which is now under development in collaboration with Philips Research in Aachen. ASR will also solve the problem of entering caller address information.

4.2 Speaker Verification

Mail ordering is a transaction process that requires a certain amount of protection against fraud. In the present version fraud protection hinges on the PIN code of the user's Scopecard. It is well known that this code is easy to steal. Thus, we intend to introduce speaker verification techniques to reduce fraud. This will be done in close collaboration with Card Services, taking advantage of the fact that much work on the coupling of databases has already been done for allowing billing via the Scopecard.

4.3 Calling Line Identification

PTT Telecom will introduce CLI in the network in the near future. Our service will use CLI data as an extra level of security: if the caller's name and/or address do not match the database information a special verification dialogue will be entered. The interaction with residential callers will be simplified by verifying database address (and billing) information, instead of asking the caller to provide that information from scratch.

V. CONCLUSIONS

This paper has described the concepts of a consumer information server and has outlined the possibilities of such systems. The addition of facsimile and other output facilities to interactive voice response services enables a variety of new applications that require high density information to be send to the consumer.
Careful attention to the design and style of voice response menus will help to increase the use of these services, make it easier and quicker to develop new services and provide greater customer satisfaction.
Future enhancements like speech recognition and speaker verification will make more applications possible and will help to increase the quality of available services.

REFERENCES

1 Remmers, W.M.: 'The third revolution in publishing', PTT Research, September 1993, Leidschendam.

2 Hunter, R., Dixon, I. and Ablard, W.: 'Fax-on-demand systems for business and home applications', BT Technol J, Vol. 12, No 1, pp34-43 (January 1994).

---

1 The Scopecard is useful to telephone abroad and receive the bill later through an national bank account.